



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/818,172	03/27/2001	Eliot M. Case	1814 (USW 0617 PUS)	2380

22193 7590 03/02/2005

QWEST COMMUNICATIONS INTERNATIONAL INC  
LAW DEPT INTELLECTUAL PROPERTY GROUP  
1801 CALIFORNIA STREET, SUITE 3800  
DENVER, CO 80202

EXAMINER

JACKSON, JAKIEDA R

ART UNIT	PAPER NUMBER
----------	--------------

2655

DATE MAILED: 03/02/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	Application No. 09/818,172	Applicant(s) CASE, ELIOT M.	
	Examiner Jakieda R Jackson	Art Unit 2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) ☒ Responsive to communication(s) filed on 24 September 2004.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) ☒ Claim(s) 1-11 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-11 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |   |   |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)  | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date <u>July 8, 2004</u> . | 6) <input type="checkbox"/> Other: _____  |

## **DETAILED ACTION**

### ***Response to Amendment***

1. In response to the Office Action mailed June 15, 2004, applicant filed a Notice of Appeal on September 17, 2004, in which the applicant argues whether the sole independent claim 1 and the dependent claim 2 is anticipated by the prior art used in the rejection.

### ***Response to Arguments***

2. Applicant's arguments, see pages 4-6, filed September 24, 2004, with respect to claims 1 and 2 have been fully considered and are persuasive. The finality of rejection has been withdrawn. However, upon further consideration, a new ground(s) of rejection is made in view of Cecys and Parthasarathy et al.

### ***Specification***

3. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

***Claim Rejections - 35 USC § 102***

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. **Claim 1** is rejected under 35 U.S.C. 102(b) as being anticipated by Cecys (USPN 5,704,007).

Regarding **claim 1**, Cecys discloses a method for converting text to concatenated voice (figure 1, element 100 with concatenative synthesis of human speech; column 1, lines 32-42) by utilizing a digital voice library (figure 6(c), element 603 with column 10, lines 22-32 and 46-59) and a set of playback rules (acoustic processor annotates the input phonetic string with the appropriate vocal parameters; column 7, lines 57-64), the digital voice library including a plurality of speech items including words and syllables (words and syllables) and a corresponding plurality of voice recordings (recording waveform data samples) wherein each speech item corresponds to at least one available voice recording (column 1, lines 32-42 with column 7, lines 45-49), the method comprising:

training (linking human speech segments to build syllables) the digital voice library to associate each syllable speech ("Mary" is comprised of two syllables) item with a literal text syllable of the particular syllable speech item ("m-Eh" and "r-IY"; column 11, lines 4-25).

***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. **Claims 2-4** are rejected under 35 U.S.C. 103(a) as being unpatentable over Cecys in view of Parthasarathy et al (USPN 6,233,555), hereinafter referenced as Parthasarathy.

Regarding **claim 2**, Cecys discloses the method further comprising:

receiving a sequence of words (the input text string "The cat sleeps") including known words that correspond to word speech items in the digital voice library (column 7, lines 50-55 with prerecorded speech; column 1, lines 32-42); and

converting each known word into a word speech item in accordance with the digital voice library (TTS; figure 1, element 100), but lacks including unknown words;

Parthasarathy discloses a method for speaker identification using mixture discriminant analysis to develop speaker models comprising:

receiving a sequence of known and unknown words (column 6, lines 20-41), and parsing the unknown word (utterance not known) to determine a sequence of literal text syllables (\*column 3, lines 22-32) and converting the text syllable sequence (text) to a sequence of syllable speech items (phoneme conversion) in accordance with

Art Unit: 2655

the digital voice library (dictionary; column 3, lines 16-32), to allow training and enrollment of new phrases (column 1, lines 47-18).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Cecys method wherein it includes unknown words, to allow training in enrollment of new phrases, which creates and stores phone transcriptions (column 3, lines 22-34) for an accurate and reliable speaker identification system (column 1, lines 26-28).

Regarding **claim 3**, Cecys discloses the method further comprising:

converting the sequence of word speech items and syllable speech items (column 11, lines 20-25) into a sequence of voice recordings (recorded sound sample; column 8, lines 29-52) in accordance with the set of playback rules (annotating the input string with the appropriate vocal parameters; column 7, lines 45-64).

Regarding **claim 4**, Cecys discloses the method further comprising:

generating voice data (voice source) based on the sequence of voice recordings (recorded sound sample) by concatenating adjacent recordings in the sequence of voice recordings (concatenative speech synthesizer; column 8, lines 29-47 with column 1, lines 32-42).

(\*the invention is described below using phones as the primary language segmentation unit, but it may be appreciated that the invention include the use of other language segmentation units, such as syllables or acoustic sub-units, for example; column, 2, lines 45-50)

Art Unit: 2655

8. **Claim 5** is rejected under 35 U.S.C. 103(a) as being obvious over Cecys in view of Parthasarathy, as applied to claim 2, and in further view of Karalli et al. (USPN 5,668,926), hereinafter referenced as Karalli.

Regarding **claim 5**, Cecys in view of Parthasarathy discloses a method for converting text to concatenated voice, but lacks training the dictionary by "utilizing a neural network having an input and an output to train the digital voice library with the neural network receiving the literal text syllable of the particular syllable speech item as input and with the neural network outputting the associated syllable speech item".

Karalli teaches the use of neural networks to train the text-to-speech system (col. 2, lines 21-33).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Cecys in combination with Parthasarathy's method, as taught in Karalli, in order to populate the diphone dictionary in the efficient manner and also provide an effective method of resolving ambiguous inputs to the dictionary.

9. **Claim 6** is rejected under 35 U.S.C. 103(a) as being obvious over Cecys in view of Parthasarathy, as applied to claim 2, and in further view of Walker (USPN 6,510,413).

Regarding **claim 6**, Cecys in view of Parthasarathy discloses a method for converting text to concatenated voice, but lacks training the digital library by "manually associating each syllable speech item with the literal text syllable of the particular syllable speech item".

The process of manually populating any look-up table (or dictionary) is similar to the process of inserting the words in a foreign dictionary (For example, English-Spanish). In that case, an editor/writer manually creates a mapping between each English word and its Spanish translation. Alternatively, similar mappings are using in computer ads. For example, "hosts" file on Windows operating system allows the user to manually enter the mappings between the IP addresses and host names. Other examples in the computer arts abound (such as address books). Therefore, manually adding entries to tables/dictionaries of various information is by no means an original concept and is well-known in many arts, including computer hardware and software.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Cecys in combination with Parthasarathy method to manually associate each literal text syllable with the corresponding syllable speech item since this would be the most straightforward and "brute force" method of training the dictionary.



10. **Claims 7-10** are rejected under 35 U.S.C. 103(a) as being obvious over Cecys in view of Parthasarathy, as applied to claim 2, and in further view of Lin et al. (USPN 6,076,060), hereinafter referenced as Lin.

Regarding **claim 7**, Cecys in view of Parthasarathy discloses a method for converting text to concatenated voice, but lacks "parsing the unknown word to determine a sequence of literal text syllables and known words, and converting the sequence to a sequence of syllable speech items and word speech items in accordance with the digital voice library."

Lin teaches parsing the unknown word into a sequence of syllables and word speech items (col. 6, lines 56-60) that are later converted to speech sounds (fig. 2, element 16).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Cecys in combination with Parthasarathy method, as taught in Lin, in order to eventually create a dyphone representation of each unknown word so it could be synthesized by speech synthesizer that requires an input of dyphones to produce the output sound.

Regarding **claim 8**, Cecys in view of Parthasarathy discloses a method for converting text to concatenated voice, but lacks the method comprising:

parsing the unknown word in the forward direction to determine any known words;

parsing the unknown word in the reverse direction to determine any known words where any known words overlap, selecting the larger word;

parsing the unknown word in the forward direction to determine any literal text syllables; and

parsing the unknown word in the reverse direction to determine any literal text syllables.

Lin et al. teach parsing the words in from left-to-right and from right-to-left in order to determine sub-words and literal text symbols (col. 3, lines 45-53). Also, the large words are chosen first (col. 3, lines 55-58).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Cecys in combination with Parthasarathy method, as taught in Lin, in order to create an efficient parsing technique that more closely matches the way words are parsed when spoken by humans. This method of parsing is less likely to miss important sub-stings in unknown words.

Regarding **claims 9 and 10**, Cecys discloses the use of different voice sources having different voice colorations. The acoustic processor annotates the input phonetic string with the appropriate vocal parameters, so as to inform the Speech Synthesizer which, and how much of each, voice source to use for each phonetic element.

11. **Claim 11** is rejected under 35 U.S.C. 103(a) as being obvious over Cecys in view of Parthasarathy, as applied to claim 2, and in further view of Carter et al. (USPN 6,600,814), hereinafter referenced as Carter.

Regarding **claim 11**, Cecys in view of Parthasarathy discloses a method for converting text to concatenated voice, but lacks "for each unknown word, after the unknown word is parsed, storing results of the parsing in the digital voice library so that a next encounter with the same unknown word may be handled more efficiently."

Carter teaches storing processed portions of text in the text-to-speech system to alleviate the load on the system (col 2, lines 30-39).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Cecys in combination with Parthasarathy method, as taught by Carter, to store the parsed results of unknown words so that next attempts with the same words were handled more efficiently. This concept of "caching" data for future reference is extremely well-known and widely used in the art of computing.


**Conclusion**

12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jakieda R Jackson whose telephone number is 703.305.5593. The examiner can normally be reached on Monday through Friday from 7:30 a.m. to 5:00p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on 703. 305.4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

JRJ  
February 23, 2005

  
David Ometz  
Primary Examiner